Project Title: Climate Change Research in Support of Hawaiian Ecosystem Management: An Integrated Approach (#G12AC20502)

Data Input – New Collection (one form per input collection)

1	None
Description:*	
Budget:*	
Format:	
Data Processing & Scientific Workflows:	
Protocols:*	
Backup & Storage:	
QA/QC:*	
Metadata:	
Volume Estimate:	
Archive Orgs:	
Access & Sharing:	
Exclusive Use:*	
Restrictions:*	
Citation:	
Digital Object Identifier (DOI)/Link:	
Contact:*	

*Indicates that this information was requested as part of the PDMP.

Data Input – Existing Collection (one form per input collection)

1	Best filled Monthly Mean Rainfall Station Data for the Hawaiian Islands
Description:*	Monthly precipitation data sets that have been developed for the Rainfall Atlas of Hawaii project (Lead-PI T. W. Giambelluca). It contains 1104 stations with mostly complete monthly precipitation amounts 1920-2010.
Format:	Spreadsheet tables (MS Excel, CSV-format)
Source:	http://rainfall.geography.hawaii.edu/
Data Processing & Scientific Workflows:	Use as acquired
Backup & Storage:	The original input data were copied to a local disc on a server within the IPRC and backed up on second hard-drive on a local PC.
Volume Estimate:	30MB
Access & Sharing:	NA
Restrictions:*	Unrestricted use. Proper reference to source required.
Fees:	None
Citation:	Giambelluca, T.W., Q. Chen, A.G. Frazier, J.P. Price, Y.-L. Chen, P.-S. Chu, J.K. Eischeid, and D.M. Delparte, 2012: Online Rainfall Atlas of Hawaiʻi. *Bull. Amer. Meteor. Soc.*, doi: 10.1175/BAMS-D-11-00228.1

*Indicates that this information was requested as part of the PDMP.

2	NCDC daily-mean meteorological station data for Hawaii
Description:*	Daily mean temperature and daily precipitation amount data sets from the NCDC's Global Historical Climatology Network (GHCN-Daily).
Format:	ASCII text files in GHCN specific format.
Source:	http://www.ncdc.noaa.gov/oa/climate/ghcn-daily/
Data Processing & Scientific Workflows:	We will download the latest version data set and extract temperature and precipitation data sets for the Hawaiian Island stations covering 1950-2011. Preprocessing will include reformatting to continuous time series and more compact binary data stream files.
Backup & Storage:	RAID discs and USB and/or DVD for additional backup.
Volume Estimate:	1GB
Access & Sharing:	NA
Restrictions:*	Unrestricted use. Proper reference to source required.
Fees:	None
Citation:	Menne, Matthew J., Imke Durre, Russell S. Vose, Byron E. Gleason, Tamara G. Houston, 2012: An Overview of the Global Historical Climatology Network-Daily Database. J. Atmos. Oceanic Technol., 29, 897–910. doi: http://dx.doi.org/10.1175/JTECH-D-11-00103.1

*Indicates that this information was requested as part of the PDMP.

Data Input – Existing Collection (one form per input collection)

3	CMIP5 climate change model scenarios
Description:*	A collection of the latest Coupled Model Intercomparison Project Phase 5 (CMIP5) climate change scenarios including the historical runs for the mid-19th to early 21th century. Monthly data were downloaded from the official data servers (all existing model outputs) for a selected number of global atmospheric data from the CMIP5 database and a locally installed archive at the IPRC. Daily data will be obtained at a later stage for a representative subset of the full model collection.
Format:	NetCDF (CF)
Source:	http://pcmdi9.llnl.gov/esgf-web-fe/
Data Processing & Scientific Workflows:	Unix/Linux based scripts to download data sets from CMIP5 Application of Climate Data Operators (CDO) to extract variables on single atmospheric levels; merging of time-segments files into a single continuous time series files.
Backup & Storage:	Currently, the extracted and merged data files are stored on a local USB disc. We will transfer the project-related input data files to RAID discs.
Volume Estimate:	6TB
Access & Sharing:	NA
Restrictions:*	Unrestricted use. Proper reference to source required.
Fees:	None
Citation:	Taylor, K.E., R.J. Stouffer, G.A. Meehl: An Overview of CMIP5 and the experiment design." Bull. Amer. Meteor. Soc., 93, 485-498, doi:10.1175/BAMS-D-11-00094.1, 2012.

*Indicates that this information was requested as part of the PDMP.

Software and Other Needs (one form per software package)

1	Climate Data Operators
Description:*	Climate Data Operators (CDO) is Unix/Linux based program to read and manipulate NetCDF climate data sets, which is the standard in climate modeling.
Restrictions:*	This software is public domain
Fees:*	Free of charge
Source/Link:	https://code.zmaw.de/projects/cdo

2	NetCDF library
Description:*	The NetCDF library supports programming languages such as we use here (Fortran77, Fortran90) to read and write NetCDF files. NetCDF is the standard model output format in the climate modeling community. CMIP5 model data are in NetCDF format.
Restrictions:*	This software is public domain
Fees:*	Free of charge
Source/Link:	http://www.unidata.ucar.edu/software/netcdf/

3	UDUNITS
Description:*	UDUNITS is another support library for the NetCDF data files. The NetCDF library depends on UNUNTS.
Restrictions:*	This software is public domain
Fees:*	Free of charge
Source/Link:	http://www.unidata.ucar.edu/software/udunits/

4	Fortran Compiler ifort and gfortran
Description:*	We use a number of specific programs written in Fortran77 and Fortran90. Compilers translate the code into executable programs.
Restrictions:*	'ifort' is available for IPRC computers via a shared license. 'gfortran' is public domain software available for most Unix/Linux and MAC-OS systems.
Fees:*	'intel' license provided at no extra cost through the IPRC. 'gfortran' is free of charge
Source/Link:	http://software.intel.com/en-us/intel-compilers

Software and Other Needs (one form per software package)

5	R
Description:*	R is an open-source statistical software package similar to S-plus. A number scientific users and a dedicated developer community make use of and develop R. Standard and state-of-the-art statistical analysis algorithms and packages are available for R, and many of our statistical downscaling procedures are computed in R.
Restrictions:*	none
Fees:*	Free of charge
Source/Link:	http://cran.r-project.org/

6	GMT
Description:*	Generic Mapping Tools (GMT) is a collection of data processing and visualization tools for geospatial data. We use it for producing maps for our geospatial data sets.
Restrictions:*	none
Fees:*	Free of charge
Source/Link:	http://gmt.soest.hawaii.edu/

7	python
Description:*	We use python as a scripting language to control and automate data preprocessing, data analysis methods, visualization procedures.
Restrictions:*	none
Fees:*	Free of charge
Source/Link:	http://www.python.org/

*Indicates that this information was requested as part of the PDMP.

Data Output (e.g., Project Data Product) (one form per output)

1	[Drought and heat wave statistics]
Description:*	We will produce data table summaries for the statistically downscaled changes in the recurrence time of droughts and heat waves.
Budget:*	Approx 30% of total budget $86,000 (including all process steps)
Format:*	Standard spreadsheet tables for Excel and similar programs (XLS, CSV-format).
Data Processing & Scientific Workflows:	Daily and monthly station data of rainfall will be converted into drought-probabilities, from which the recurrence time can be estimated for the observed past decades. Statistical downscaling methods will link the station-based drought probability indices to large-scale climate modes. Future climate scenarios from CMIP5 will be analyzed for the mid and late 21st century and estimates for changes in drought frequencies will be obtained via statistical downscaling. Similarly, temperature time series will be converted into heat-wave indices and changes in the recurrence of heat waves will be estimated.
Quality Checks:*	Standard quality checks will be performed during the data processing by introducing test datasets with exactly known statistical properties into the actual data sets. Furthermore, we will follow cross-checking of the consistency between input and output data sets after each analysis/processing step.
Backup & Storage:	Data and developed computer program source codes and scripts will be stored on the RAID discs and backups of the developed and tested programming codes and scripts will be stored on independent USB discs/hard discs.
Metadata:	TBD
Volume Estimate:	10-100 GB (depending on the backup level of the intermediate data products generated through the analysis process) output products itself 10-20 MB.
Access & Sharing:	Data will be shared with partners in related PICCC / PICSC projects
Exclusive Use:*	We request a maximum of a two year extension following the conclusion of the funding period to write manuscripts and publish our results. During this time data will be shared upon request with non-project related researchers and data and project managers. After the extension period, no restrictions to the data access.
Archive Orgs:	We will provide the data and metadata also to NCCWSC for archiving purposes. The IPRC's data center APDRC is suggested as host for the data products (online) with user-friendly interfaces (http://apdrc.soest.hawaii.edu/).
Restrictions:*	We do not anticipate restrictive use of the data after publication.
Citation:	TBD
Digital Object Identifier (DOI)/Link:	TBD
Contact:*	TBD

*Indicates that this information was requested as part of the PDMP.

Data Output (e.g., Project Data Product) (one form per output)

1	[Statistically downscaled maps of seasonal mean precipitation changes]
Description:*	We produce updates for the seasonal mean precipitation changes projected for the 21st century. Spatial interpolation will be applied to create maps for the Hawaiian Island showing estimated precipitation changes and confidence ranges.
Budget:*	Approx 30% of total budget $86,000
Format:*	Shape files for GIS applications and gridded data files (NetCDF, ASCII tables), Image files (JPG).
Data Processing & Scientific Workflows:	We will follow and update the statistical downscaling methods, which were developed in previous projects (e.g. Timm and Diaz, J. Climate, **22**(16),4261–4280, doi: 10.1175/2009JCLI2833.1) , and apply it to CMIP5 model scenarios onto rainfall changes for the Rainfall-Atlas-of-Hawaii station network.
Quality Checks:*	Standard quality checks will be performed during the data processing by introducing test datasets with exactly known statistical properties into the process stream. Furthermore, we will follow cross-checking for the consistency between input and output data sets after each analysis/processing step (file size, data format, data range).
Backup & Storage:	Data and developed computer program source codes and scripts will be stored on the RAID discs and backups of the developed and tested programming codes and scripts will be stored on independent USB discs/hard discs.
Metadata:	TBD
Volume Estimate:	10-100 GB (depending on the backup level of the intermediate data products generated through the analysis process) output files: 1-10GB
Access & Sharing:	Access to partners in related PICCC / PICSC projects.
Exclusive Use:*	We request a maximum of a two year extension following the conclusion of the funding period to write manuscripts and publish our results. During this time data will be shared upon request with non-project related researchers and data and project managers. After the extension period, no restrictions to the data access.
Archive Orgs:	We will provide the data and metadata also to NCCWSC for archiving purposes. The IPRC's data center APDRC is suggested as host for the data products (online) with user-friendly interfaces (http://apdrc.soest.hawaii.edu/).
Restrictions:*	We do not anticipate restrictive use of the data after the two-year extension period.
Citation:	TBD
Digital Object Identifier (DOI)/Link:	TBD
Contact:*	TBD

*Indicates that this information was requested as part of the PDMP.

Further remarks and clarifications:

The budget allocation for the data management plan was based on a estimations, what parts of the time and work from all participating research personnel and supervisors will be spent on the different project activities that connect input data to the final output products. The hardware and material costs that are directly related to the physical storage of the data products is significantly less.

The data management plan does further not include any indirect costs that third-party institutions such as the Asia-Pacific Data Research Center (APDRC) could charge for data management services in future.